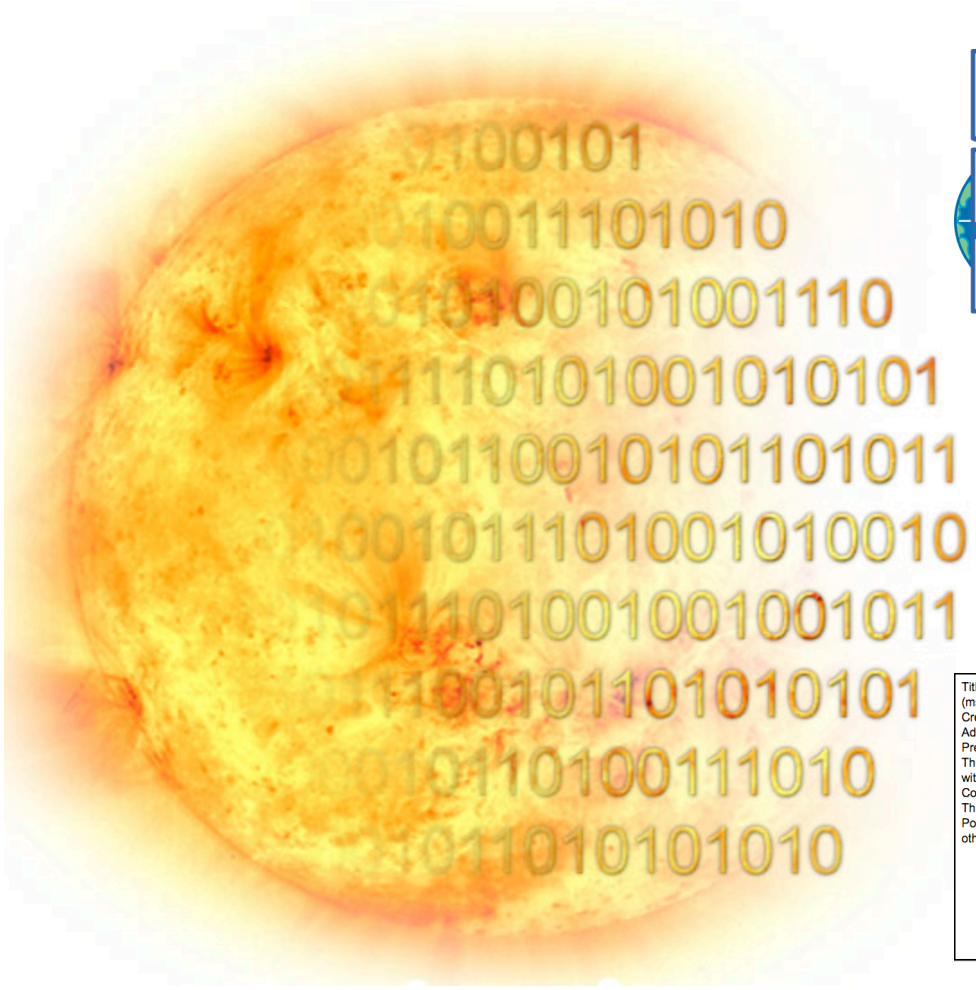
70.02

Implementation of the Virtual Solar Observatory

A.R. Davey (SwRI), R.S. Bogart (Stanford Univ.), J.B. Gurman (NASA-GSFC),
F. Hill (NSO), J. Hourclé (NASA-GSFC), P.C. Martens (Montana State Univ.),
I. Suarez-Sola (NSO), K.Q. Tian (Stanford Univ.), K. Yoshimura (Montana State Univ.)



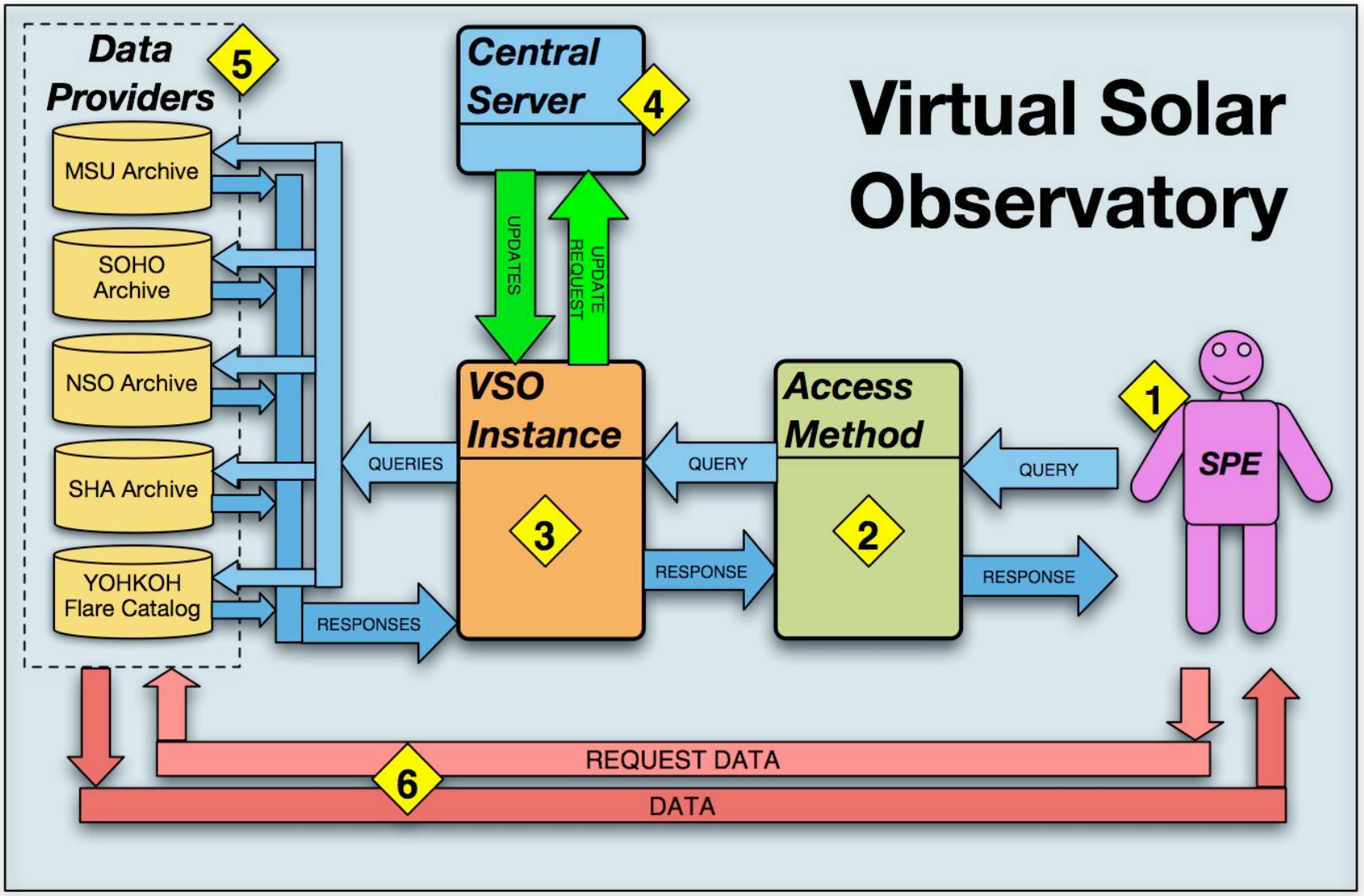
Title:
(SU_seal_201_pos.eps)
Creator:
Adobe Illustrator(TM) 7.0
Preview:
This EPS picture was not saved
with a preview included in it.
Comment:
This EPS picture will print to a
PostScript printer, but not to
other types of printers.

NSO
National Solar Observatory

Title:
(msub-M-cmyk.eps)
Creator:
Adobe Illustrator(R) 8.0
Preview:
This EPS picture was not saved
with a preview included in it.
Comment:
This EPS picture will print to a
PostScript printer, but not to
other types of printers.

Title:
(NASA Meatball Logo.eps)
Creator:
Adobe Illustrator(TM) 7.0
Preview:
This EPS picture was not saved
with a preview included in it.
Comment:
This EPS picture will print to a
PostScript printer, but not to
other types of printers.

VSO Architecture - The Standard Model (Simplified)



VSO Architecture - The Pieces

The VSO is distributed data service, providing a uniform virtualized query interface to distributed providers of solar data. Some of the pieces of the VSO are described below.

1. The User
   The user currently accesses the VSO via a browser. In future this may be a standalone program. The VSO is platform/OS and browser blind!
2. Access Method
   The primary access method to the VSO functionality is via cgi scripts currently available at Stanford and NSO. The API is being developed for integration with languages other than PERL, including a prototype JAVA interface. A laptop, standalone version of the VSO has already been demonstrated.
3. VSO Instance
   This is the VSO core that handles the user queries, decides which sites to query for data, and returns results.
4. Central Server
   The central server is used to log searches for reproducibility and attribution purposes. It also keeps an up-to-date record of the various data providers, including information on their availability. It is important to note that the availability of the central server is not required for VSO functionality.

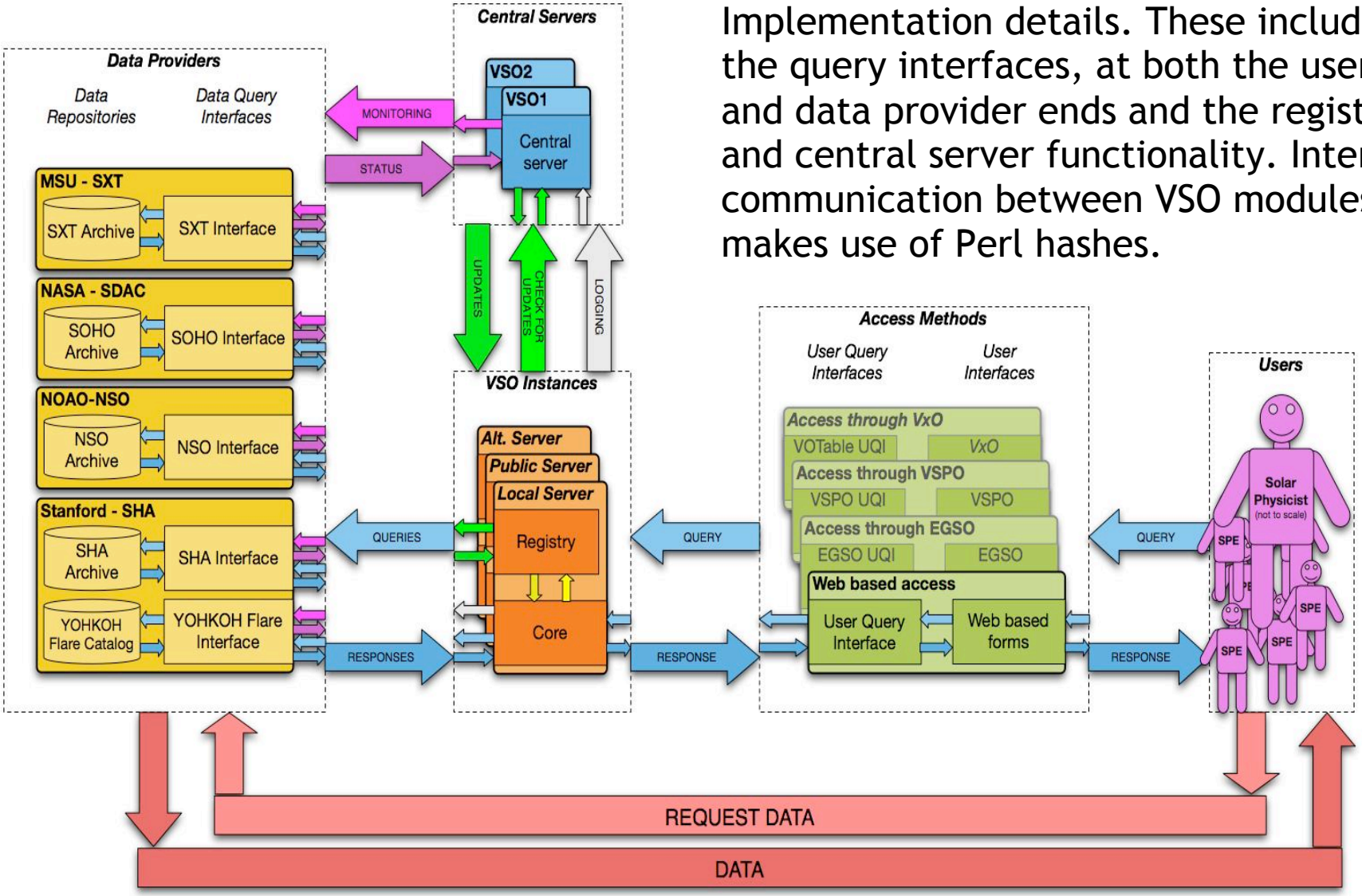
5. Data Provider

   The data providers are where the actual data searches are performed. A complete list of data providers is given below.

6. Data Request / Return

   Data can be retrieved directly via the facilities a data provider offers, or by using an intelligent VSO agent which interrogates providers regarding data availability, integrates the results and then emails you the results detailing where and when data is available.

VSO Architecture - In Detail



This image includes more of the VSO Implementation details. These include the query interfaces, at both the user and data provider ends and the registry and central server functionality. Internal communication between VSO modules makes use of Perl hashes.

VSO Architecture - SOAP and Perl

The VSO core is implemented in the Perl language. Perl was chosen because of the prevalence of web services already written in the language and the availability of already developed packages for example, for networking, database interaction and lexical matching. The VSO makes use of the Simple Object Access Protocol (SOAP), using the Perl SOAP::Lite module, to provide platform and OS independent access to Solar data. SOAP is the networking glue that allows the VSO to talk to data providers and integrate with other virtual observatory projects such as EGSO and COSEC.

API

External access to the VSO is via the published API. Details and example code are available from http://vso.stanford.edu/0.6/api/index.html. Both Perl and JAVA access is demonstrated and a VSO WDSL is also available.

Available Data from the VSO

VSO data is currently served from MSU, NSO, SDAC and Stanford.

MSU:

Yohkoh – SXT, HXT, WBS and BCS.

NSO:

Evans – Spectroheliograph.
KPVT – Spectromagnetograph, 512-channel Magnetograph.
McMath – Solar FTS Spectrometer.
SOLIS – VSM.

SDAC:

SOHO – EIT, MDI, CELIAS, CDS, COSTEP, ERNE, GOLF, LASCO, SUMER, SWAN, UVCS, VIRGO.

Stanford:

GONG – Big Bear, Udaipur, Mauna Loa, El Teide, Learmonth, Cerro Tololo.
Mt. Wilson – MOF/60.
SOHO – MDI.
TON – Big Bear, Tenerife.

OVRO:

OVRO – OVSA (currently via Stanford proxy).

A number of catalog resources, NGDC data and index data from the Royal Observatory of Belgium are currently being developed for inclusion.

Currently Available Searches

The latest VSO search capabilities may be tested at:

http://vso.stanford.edu/ and http://vso.nso.edu/0.6

The following search types are available:

Search All By Time

A very simple form that automatically searches all datasets registered by all providers in the selected time interval.

Search Selected Instruments By Time

Performs a joint time-based query on data from selected instruments or archives only.

Search Selected Observables By Time

Searches all data sets for data in the selected time interval matching selected criteria for physical observable.

Select Observables, Instruments and Time

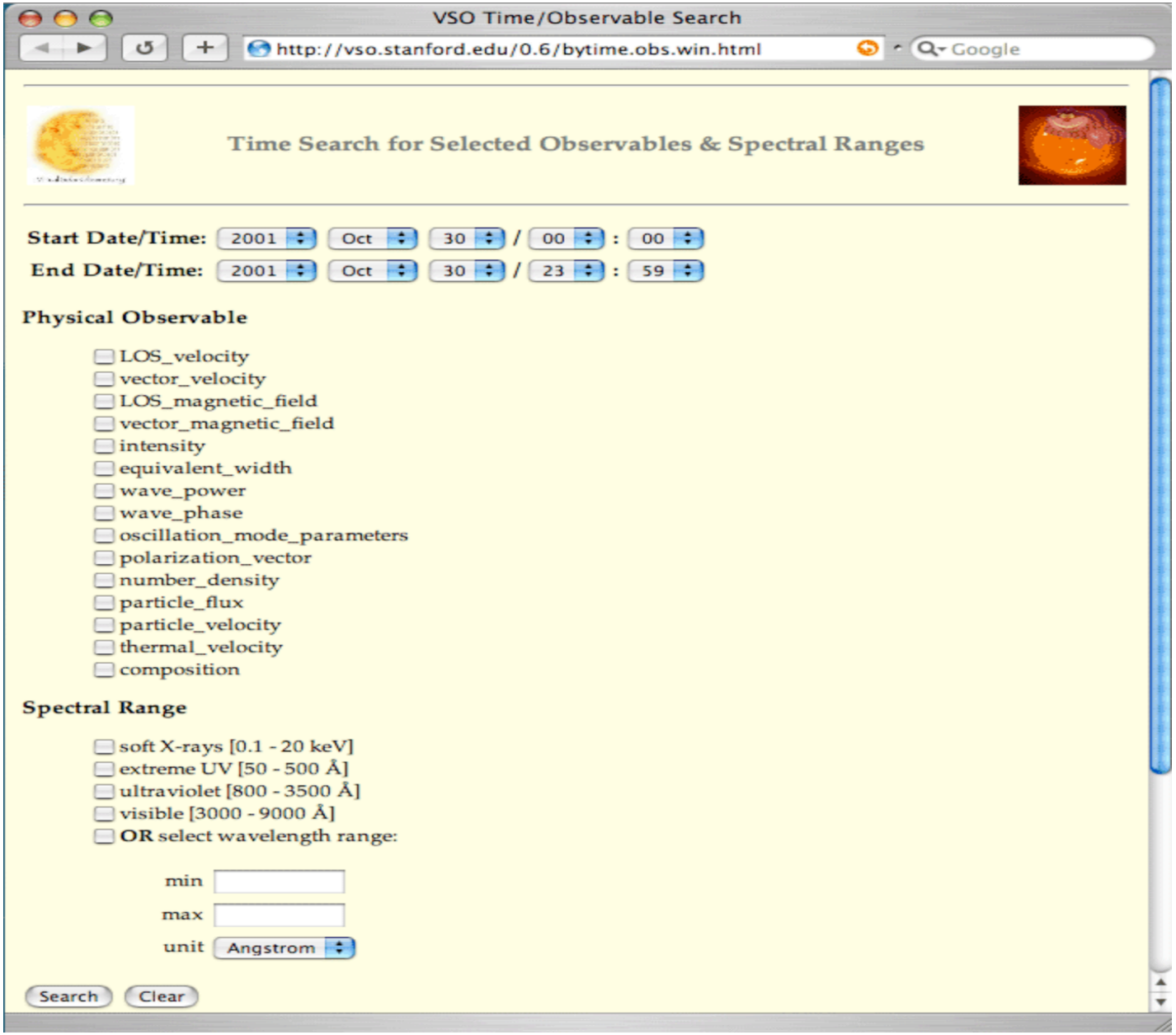
Searches selected instrument or archive data sets for data in the selected time interval matching selected criteria for physical observable.

Search Selected Observable / Spectral Range By Time

Searches all data sets for data in the selected time interval matching selected criteria for physical observable and/or spectral range.

Sample VSO Query Interface

The observable / spectral range / time query interface is shown below.



VSO Data Model

The VSO has developed a data model via which we describe the datasets held at a particular site. This data model is used to determine which sites are to be contacted for a given query. We also make use of common names used by Solar Physicists to describe various datasets, to make both the data descriptions and interfaces more intuitive. Whilst this data model has been developed independently we will be working with the EGSO to try and unify our data descriptions.

Element Group	**V1.7 Element Name or** FITS Keyword	**Sample Values/Explanation**
1. Observing Time	**Observation_Time** T_OBS	*time* The time at which the data comprising an atomic data set were originally recorded. If the duration of the data in the atomic data unit is large compared with the search time resolution, the Observation_Time is to be understood to correspond to the *center* (mid-point) of the observation(s), weighted as appropriate. For purposes of the Data Model, Observation_Time is given in calendar-clock form, *e.g.*2004.03.08_16:25. Times are assumed to be UTC. The time resolution is one minute, so for much data the conversion from say start time of an exposure to Observing_Time should not matter. Likewise the conversions between UTC and other units such as ET, TAI, and GPS should not be a matter of much concern. A data match is assumed to include all data from 30 seconds before the target time to 30 seconds after, inclusive (closed at both ends), so that a data Observation_Time can in principle fall into two adjacent target times. Note that since Jan 1, 1999, TAI = UTC + 32 sec, and GPS = UTC + 13 sec. *number* The interval, in seconds, between the start and end of observation in the atomic data unit. For a single image or spectrum, this is simply the exposure time; for a movie, it is the time difference between the start of the first image and the end of the last. *Nnumber* The interval, in seconds, between successive time samples
	Duration T_LENGTH	

	Time_Step T_STEP	
2. Observer Location	TBD	*TBD* - (Note – only important for in-situ measurements. Actual set of elements to be obtained from VSPO)
3. Spectral Sampling	**Wave_Type** WV_TYPE **Wave_Bands** WV_NBAND **Wave_Minimum** WV_MN_xx **Wave_Maximum** WV_MX_xx **Wave_Step** WV_ST_xx	*string* The type of spectral data. Allowed values: "broadband" or "narrowband" *number* The number of wavelength bands in the observation (assumed value of 1 if not specified) *number* Minimum wavelength for band xx, in terms of **WV_UNIT.** Used when **WV_TYPE** = "broadband" *number* Maximum wavelength for band xx, in terms of **WV_UNIT.** Used when **WV_TYPE** = "broadband" *number* Dispersion, wavelength units per pixel for band xx, in terms of **WV_UNIT.** Used when **WV_TYPE** = "broadband"
4. Global Acoustic Mode Sampling	**Degree_Minimum** L_MIN **Degree_Maximum** L_MAX **Degree_Step** L_STP	*number* Minimum value of spherical harmonic degree _ *number* Maximum value of spherical harmonic degree _ *number* Spacing between spherical harmonic degree _
5. Observable	**Physical_Observable** PHYS_OBS	*string or menu* An identification of the physical observable(s) represented by the data set. Current allowed values: "LOS_velocity", "vector_velocity" "LOS_magnetic_field", "vector_magnetic_field, "intensity", "equivalent_width", "wave_power", "wave_phase", "oscillation_mode_parameters", "polarization_vector", "number_density", "particle_flux", "particle_velocity", "thermal_velocity", "composition"
6. Spatial Location	**Observation_Center_West**	*numbers* A pair of coordinates specifying the location of the center of the image data

	Observation_Center_North CENT_WST CENT_NRT	circle, in arc-seconds, with respect to the Earth-Sun line at the nominal Observation_Time. This origin is close to the center of the apparent solar image for Earth-based or near-Earth observers, but not necessarily so for deep space observations. The North coordinate is measured in the direction of the Carrington axis (RA 286°.13, _ 63°.87 J2000.0), and the West coordinate in the direction of solar rotation. *number* The radius of the bounding circle, in arc-seconds. For the VSO data model the bounding circle is to be understood as either the maximum inscribed circle in the bounding data rectangle (polygon), or the minimum circumscribed circle, dpeending on whether the query is for included data (presumably the normal default) or excluded data, respectively
	Bounding_Radius R_BOUND	
7. Data Source	**Observatory** OBSERVTY	*string* Name of specific ground-based observatory or spacecraft that obtained the data. Examples: “SoHO”, “SDO”, “Yohkoh”, “TRACE”, “Kitt_Peak”, “BBSO”, “Sac_Peak”, “Mees”, “Wilcox”, “Mt_Wilson”, “GONG”, GONG_Big_Bear” *string* Name of specific instrument that obtained the data. Examples: “SUMER”, “SX-T”, “MDI”, “VSM”, “LASCO_C1”, “LASCO_C2”, “ASP”, “NIM”, etc.
	Instrument INSTRUMT	

OVRO Registry Entry

```
<xml>
<version>0.6</version>
<provider>OVRO</provider>
<contact>Dale Gary</contact>
<uri>http://l5-m5.stanford.edu/VSO/OVROi</uri>
<proxy>http://l5-m5.stanford.edu/cgi-bin/ovro/ovroi.cgi</proxy>
<available>1</available>
<dataset>
      <source>OVRO</source>
      <instrument>OVSA</instrument>
      <physobs>intensity</physobs>
      <wave>
            <wavemin>1.2</wavemin>
            <wavemax>18</wavemax>
            <waveunit>GHz</waveunit>
      </wave>
      <time>
            <start>20000316162315</start>
      </time>
</dataset>
</xml>
```

Implementation Kit

We are in the process of developing an implementation kit for data providers. This kit will contain several elements including comprehensive documentation.

A web form that prospective data providers will use to describe the data holdings and create a registry entry.

Sample code, demonstrating the data provider query interface and interaction with for example, ASCII files and MySQL databases.

Programming expertise! Effort to help you bring your data online is available.

The Future!

In the next couple of months we will be:

- Adding more data providers including NGDC and MLSO.
- Incorporating catalogues and integrating them into the search capabilities.
- Implementing the central logging service for queries.
- Implementing registry updates and availability.
- Finalizing the API.
- Finishing the Data Providers' Implementation Kit.
- Releasing VSO Version 1.0 – Date TBD!